

***SOAP rings the globe**

Robert S. Bauer
LaTrobe University
Melbourne, Australia

A recently published article on the Chaozhou dialect of Min spoken in Singapore states that the word *sak⁷ buŋ²* 'soap' (*sap⁷ bun²* in other Min dialects of China and Taiwan) is a borrowing of Malay *sabun* (Li Yongming 1991:58). This claim may very well be true as far as it goes. But in directing our attention to only one small corner of the globe, it overlooks the interesting position occupied by the Min and Malay terms within the much larger and very colorful global mosaic to which Min and Malay have contributed only a few of the pieces. A close look at the lexeme 'soap' in languages around the world reveals that many languages have inherited or borrowed the same word, viz., many of the modern forms meaning 'soap' are ultimately derived from the same etymon, Proto-Indo-European **soib-on* 'soap'. This remarkable transmission of 'soap' into the world's languages has now reached global proportions, so that it seems to be one of the world's most widely borrowed words of material culture.

Cultural transmission is one of several mechanisms for introducing new words into a language. When a society adopts a new cultural object and an activity associated with it, the language of the borrowing society may follow one of two processes (or even both) to name the new object and activity, i.e., phonetic transliteration or semantic translation. In the first process the borrowing language uses its own phonetic resources to transliterate the name of the new object and activity as they are pronounced in the donor language. Japanese well illustrates this process. On the inside back cover of a Japanese-English dictionary for Japanese students of English (Nakajima 1985), there is a colorful page labelled 'Games and Sports' with cartoon-like figures of boys and girls engaged in various activities depicting nine kinds of games and sports. Beneath the pictures the names of these games and sports have been transcribed in both Japanese (in *kana* and *kanyō*) and English; the Japanese terms can be divided into the following two groups on the basis of their origin:

1. *English loanwords:*

- (1) saikuringu < "cycling"
 (2) sukēto bōdo < "skateboard"
 (3) furizubī < "frisbee"
 (4) sāfuin < "surfing"

2. *Japanese words*

- (1) henka = "fight"
 (2) sakanatsuri = "fishing"
 (3) navatobi = "jump rope"
 (4) onigokko = "tag"
 (5) umatobi = "leapfrog"¹

The group of English loanwords comprises the transliterated names of two objects and two activities associated with objects which have been adopted from English-speaking, American culture (aside from the question of who invented the bicycle, skateboard, frisbee, surfboard, etc., America is probably responsible for the commercial development of these various instruments of sport). The group of Japanese words includes two activities performed with instruments, i.e., fishing line and rope, which probably have a very long history and an independent origin within Japanese culture.

Standard Chinese, on the other hand, has generally preferred to use its own lexical resources to translate semantically the names of new inventions borrowed from the West; cf. the following two examples:

huǒ "fire" + chē "cart" → huǒchē "steam-powered locomotive, train"

chū "out" + zū "rent" + qìchē "car" → chūzūqìchē "car for renting out" = "taxi"

In contrast, Hong Kong Cantonese, which has been in intimate contact with English over the past 150 years, more resembles Japanese by favoring the phonetic transliteration of English loanwords; cf. tīksí "taxi", fēilám "film", tsy̌kūlīk "chocolate".

Lexical borrowing via phonetic transliteration is a common, pervasive phenomenon among languages in contact. I have been intrigued by the possibility that a loanword could be passed along from one language to another so that eventually it has girdled the globe. On the basis of both cognate and loan relationships between an etymon and its modern reflexes, which etymon has the greatest number of modern reflexes in the world's languages? If we consider all modern reflexes related through cognation and borrowing as being the "same" word, the question can be rephrased as follows: What is the world's most widely-transmitted etymon? Which etymon displays the most widespread geographical distribution across the world's languages? Finally, can a loanword's transmission among borrowing

¹ Literally "leap-horse". [Ed.]

languages exhibit global dimensions? I think the answer to the first three questions is Proto-Indo-European **soib-on* "soap", and the answer to the last question is "yes" with cognate and loan reflexes of **soib-on* demonstrating how this is so.

This discussion begins with words for soap in the Min dialects of Chinese which are phonetically quite different from those in other neighboring Chinese dialects, and therefore set themselves apart as a unique group. Several publications on the Min dialects have attributed the origin of the Min forms to the borrowing of Malay *sabun* "soap", e.g. Li Yongming 1991:58; *Hanyu Fangyun Cihui* (Lexicon of Chinese Dialects), Beijing University 1964:153; and Yuan 1983:299. There is no question that the Min and Malay forms are so phonetically similar that they must be related in some way. However, as will be described below, the etymological story of soap has attained global dimensions; there are many more links in the linguistic chain besides this one between Min and Malay.²

The Appendix (*Global Distribution of Modern Reflexes for SOAP*) indicates that the phonetic shape *sabun* which is phonetically similar to forms in the Min dialects appears in a number of genetically related and unrelated languages that encompass the globe. The modern forms for soap listed in the Appendix can be classified under four main phonetic shapes, viz., *sabun*₁, *sabun*, *sabu*, *sop* (and their variants). One etymon underlies these modern reflexes in the languages of the Indo-European, Uralic, Semitic, Niger-Congo, Dravidian, Altaic, Sino-Tibetan, Austroasiatic, Austronesian, Kadai, Amerind, and Indo-Pacific families. According to Buck (1949:453), the etymon of "soap" is Germanic **saip(1)ōn*; he claims the ancient Greeks and Romans did not possess soap but borrowed it and its name from Germanic tribes. Watkins (1985:56) defines the Germanic root as "dripping thing, resin" and states that it referred to "a reddish hair dye used by Germanic warriors to give [themselves] a frightening appearance." From this Germanic root Watkins has reconstructed Proto-Indo-European **soib-on*.

Similar forms for "soap" have passed into a number of languages distributed across Europe, Africa, the Middle East, South, East, and Southeast Asia, the Pacific including Australia and Polynesia, and North America. The precise details associated with the adoption of the word for soap by so many languages must be quite complex; but it is reasonable to believe that the mode of earliest lexical transmission was via trade contacts

² In addition to "soap", the lexemes "chocolate" and "coffee" may also belong to the pantheon of loanwords which have achieved the status of global distribution.

among Europeans, Arabs, Indians, and Asians. Some of the forms from languages distributed across Asia may be related to the arrival into this area of early European sailor-explorers, probably Portuguese, who carried soap with them and introduced it to the peoples with whom they came in contact. Portugal's colonial presence in Asia began in the early 16th century, with the establishment of Goa on the west coast of India in 1510 and Macao on the southeast coast of China in 1557. Portugal's first diplomat to Siam arrived in 1511, and a contingent of 120 Portuguese mercenaries were employed as bodyguards and military advisers by King Chairacha of Ayudhya (ruled 1534-1447) (Wyatt 1984:88-9).

In attempting to identify the donor languages from which the loanwords were borrowed, we need to keep in mind the varying phonetic shapes of words in the possible donor and borrowing languages. Our assumption is that the phonetic shape of the loanword will be as similar as possible to that of its source. On the basis of their phonetic shapes, loanwords for "soap" in Asian languages can be divided into two types, *sabu* and *sabun*, one with an open second syllable and one with a second syllable closed by a dental nasal consonant. We can see that among the modern European languages, Portuguese *sabão* with its nasalized second syllable is phonetically closest to *sabu*(n).³ The Portuguese etymology has already been adduced for *Siamese* *sabū*: (Haas 1964:521), for *Akha* *sãbya*, and *Mpi* *sa²bu²*, two Loloish languages spoken in Thailand (Bradley 1979:332-3), and for *Lahu* *ṣā-pē* (Matisoff 1988:1168). We can also compare *Cambodian* *sa:bū*, *Lao* *sa³ bu²*, and *Dai-Dehong* *sa:u² pɛu⁶*.

The words for "soap" in Arabic, Malay, and Min terminate in a dental nasal stop consonant. *Malay* *sabun* may have been borrowed from Arabic or some Indo-Iranian language: cf. *Arabic* *sabuun*, *Nepali* *sābun*, *Hindi/Urdu* *sa:bun*, *Persian* *sābun*. Arabic in turn could have borrowed its term from Greek or a language belonging to the Romance or Indo-Iranian branches. Among the Altaic languages, we find *Turkish* *sabun*, *Uighur* *sapun*, *Kazakh* *sabən*, *Tartar* *sabən*, *Kirghiz* *samən* (with nasalization of the bilabial stop). Although Japan's first European contacts in the Japanese Islands began with Portuguese sailors who arrived in Japan in 1543 (and were followed soon after by St. Francis Xavier who landed in Kyushu in 1549 to embark upon his mission of Christianizing the local population), the phonetic shape of *Japanese* *fabon* is closer to the Malay, Arabic, and Turkish forms, which suggests one of these languages as its source.

³ For the moment, until I have learned more about Portuguese historical phonology and the historical development of nasalized vowels from earlier final nasal consonants in Proto-Romance, I am assuming that the pronunciation of the lexeme for "soap" in 16th century Portuguese was fairly similar to the modern pronunciation.

Japanese merchant-sailors already trading in Southeast Asia prior to the arrival of the Portuguese could have encountered soap and its name during their contacts with speakers of Malay and Arabic.

The Min group of Chinese dialects are spoken mainly in the southeast coastal provinces of Fujian and Taiwan. The Appendix indicates that sa⁷ bun² is attested in at least five Min dialects (Lin and Chen 1985:141, Zhang 1983:80). Min is the only Chinese dialect group possessing forms for soap phonetically similar to words in Indo-European and Indo-Iranian languages as well as Arabic and Malay.⁴ As may have happened in the case of Japanese, I suspect that borrowing by Min speakers may also have occurred in coastal or insular Southeast Asia through their contacts with traders speaking Arabic, Malay, or some Indo-Iranian language. Similar forms are widely distributed through the area: cf. *Vietnamese-Hanoi* sà fàvŋ, *Vietnamese-Nha Trang* sà bòn, *Northern Rglai* sabon, *Indonesian* sabun, *Paiwan* sabun, *Tagalog* sabón. Min apparently had already acquired its term for soap before the Portuguese arrived in southeast China, otherwise the phonetic shape would be more like the Thai and Cambodian forms. Whatever the identity of the original donor language, the Min forms derive ultimately and indirectly from PIE *so1b-on.

⁴ The Yue dialect group is the other major Chinese dialect family of southeast coastal China. Yue dialects have had a long history of contact with the Portuguese language through the enclave of Macao which Portugal established on the coast of Guangdong in 1557. If any Chinese dialect were to have fulfilled its opportunity to borrow directly from Portuguese its word for "soap", one could reasonably expect that Yue would have done so. As it turns out, however, standard Yue *kaan³* "soap" bears no relationship to the Portuguese term, and its origin must be quite different from that of the term in Arabic, Malay, Thai, and Min. In the list of lexical terms from standard Yue-Guangzhou (or Cantonese) given below, the Chinese character 梘 read with the pronunciation *kaan³* was specially devised to transcribe the Yue morphosyllable whose origin is unknown:

Yue-Guangzhou	番梘	<i>faan¹ kaan³</i>	"[foreign] soap" (Rao et al. 1981:53)
Yue-Guangzhou	梘片	<i>kaan³ phin³</i>	"soap flakes"
Yue-Guangzhou	梘粉	<i>kaan³ fan³</i>	"soap powder" (Rao et al. 1981:64)

Hu's very brief sketch of the Portuguese influence on the Cantonese dialect spoken in Macao provides no information about what lexical term is used for "soap" among Cantonese speakers there (Hu 1991:241-2).

APPENDIX

Global distribution of reflexes for SOAP

PROTO-INDO-EUROPEAN

*soib-on

Romance:

Latin (Late)

sāpō, sāpōnis

French

savon

Italian

sapone

Portuguese

sabão

Rumanian

săpun

Spanish

jabón

Proto-Germanic

*saip(i)ōn- 'dripping thing; resin'

Old English

sāpe

Danish

sæbe

Dutch

zeep

German

seife

Norwegian

såpe

Swedish

såpa

Yiddish

zeyf

Balto-Slavic:

Latvian

ziepes

Serbo-Croatian

sapun

Indo-Iranian:

Baluchi

ṣabun

Bengali

ḡaban

Hindi/Urdu

saabun

Lari

saby

Nepali

sābun

Persian

sābun

Sinhalese

sābān

Tajik-Wahan

sabun

Tajik-Sariqul

sufun

Other:

Armenian

sabon

Albanian

sapun

Breton	soavon
Greek-Demotic	sapun1
Greek-Katharevusa	sapon
Welsh	sebon
URALIC	
Finnish	saippua, saippio
Hungarian	szappan, sâppân
AFRO-ASIATIC	
<i>Cushitic:</i>	
Omoro	sa:muna
<i>Semitic:</i>	
Arabic	sʔabuun
Beni Amer (Tigré)	sa:bu:n
Hebrew	sabon
DRAVIDIAN	
Tamil	sōp
NIGER-CONGO	
Abua	sábà
Mbum (Mbang Mboum)	sábùlù
Sambaa	sabun1
Sukuma	sabun1
Swahili	sabun1
Tumbuka	sopo
ALTAIC	
Kazakh	sabən
Kirghiz	samən
Tartar	sabən
Turkish	sabun
Uighur	sopun
JAPANESE	fabon
TAI	
Dai-Dehong	sa:u ² pe <u>u</u> ⁶
Lao	sa ³ buu ²
Siamese	sa bùu

AUSTROASIATIC

Deang	s'pia lɔi h
Jeh	sɪm bɔŋ
Khmer (Cambodian)	sa:bù:
Jing (Vietnamese of China)	sa² fɔŋ²
Vietnamese-Hanoi	sà fàwŋ [xà-phòŋ]
Vietnamese-Nha Trang	sà bɔŋ

AUSTRONESIAN

Ami f	fafun
Balinese	sabun
Cebuano-Visayan	sabón
Fijian	sovu
Hawaiian	sopa
Indonesian	sabun
Isneg	sabón
Malay	sabun
Manobo	sabun, savun
Paiwan	sabun
Ponapean	so:pʷ
Roglai (Northern)	sabon
Tagalog	sabón
Yapese	sæ:p

PIDGIN

New Guinea Pidgin	sop
-------------------	-----

AUSTRALIAN

Kalkatungu (Kalkadoon)	θupu
------------------------	------

AMERIND

Papago-Pima	shavoni
-------------	---------

SINO-TIBETAN

Sinitic

Min-Quanzhou	sap⁷ bun²
Min-Singapore	sak⁷ bun²
Min-Taipei	sap⁷ bun²
Min-Xiamen	sap⁷ bun²
Min-Yongchun	sap⁷ bun²
Min-Zhangzhou	sap⁷ bun²

*Tibeto-Burman:**(Himalayish)*

Gurung

Jirel

Kaike

Khaling

Kham

Magar

Newari

Sunwar

Tamang

Thakali

Tibetan-Lhasa

sā:bā:nā:q

sā:bunq

sābun

sābun

sāban

sābun

sābwa, sābwal-

sā:bin

'sā:pun

sāpun

sa² pun²*(Lolo-Burmese)*

Achang-Longchuan

Achang-Luzi/Lianghe

Akha

Biyue-Laiyuan

Burmese-Modern

Hani-Gelanghe

Haoni-Shuigui

Hpun

Lahu (Black)

Lahu (Yellow)

Nusu-Northern

Mpi

Naxi

Yi-Yunnan

tshau² pia²tshau¹ pjau¹sa[~] bja_~tsha³ pjau²hsa[?]pyatsha¹ bjo³tsho³ pio²

shă(?) pyá

šâ-pē

sa¹ pē⁵tsho⁴ pio²sa² pu²tsho² pia²tshao⁴ pia²*(Other)*

Bai-Jianchuan

Bai-Dali

Jingpo

Karen

tshao³ pia²tsho³ pio¹sa³ pja³tsho¹ pia³

REFERENCES

- Ahmad, Mumtag. 1985. *Baluchi Glossary*. Kensington, Maryland: Dunwoody Press.
- Awo'i-hathe et al. 1977. *Northern Roglai Vocabulary*. Manila/Huntington Beach, CA: Summer Institute of Linguistics.
- Beijing University. 1964. *Hanyu Fangyan Cihui (Chinese Dialect Vocabulary)*. Beijing: Wenzhi Chubanshe.
- Bergman, Peter M. 1968. *The Concise Dictionary of 26 Languages in Simultaneous Translation*. New York: Bergman Publishers.
- bin Yusop, Haji A. R. 1984. *Malay-English, English-Malay Dictionary*. London: William Collins & Co., Ltd.
- Blake, Barry. 1992. Handout from "Features of Case" seminar (April 14). La Trobe University, Melbourne.
- Bradley, David. 1979. *Proto-Loloish*. London and Malmö: Curzon Press.
- Buck, Carl Darling. 1949. *A Dictionary of Selected Synonyms in the Principal Indo-European Languages*. Chicago: The University of Chicago Press.
- Chen Kang et al. 1986. *Gaoshanzu Yuyan Jianzhi (Paiwanyu) (Outline Grammar of the Paiwan Language of the Gaoshan People)*. Beijing: Minzu Chubanshe.
- Drizari, Nelo. 1979. *Albanian-English and English-Albanian Dictionary*. New York: Frederick Ungar Publishing Company.
- Elkins, Richard E. 1968. *Manobo-English Dictionary*. Honolulu: University of Hawaii Press.
- Goldstein, Melvyn C. 1984. *English-Tibetan Dictionary of Modern Tibetan*. Berkeley: University of California Press.

- Gradin, Dwight, and Thông Gradin. *Jeh Vocabulary*. Ed. by Patrick Cohen. Manila: Southeast Asia Publications.
- Gragg, Gene B. 1982. *Oromo Dictionary*. East Lansing, Michigan: African Studies Center, Michigan State University.
- Haas, Mary A. 1964. *Thai-English Student's Dictionary*. Stanford: Stanford University Press. Reprinted 1985.
- Hale, Austin. 1973. *Clause, Sentence, and Discourse Patterns in Selected Languages of Nepal, Part IV: Word Lists*. Kathmandu: Tribhuvan University Press.
- Hamblin, Charles. 1984. *Languages of Asia and the Pacific: A Travellers' Phrasebook*. London: Angus & Robertson.
- He Jiren and Jiang Zhuyi. 1985. *Naxiyu Jianzhi (Outline Grammar of the Naxi Language)*. Beijing: Minzu Chubanshe.
- Hino Shun'ya. 1978. *The Classified Vocabulary of the Mbum Language in Mbang Mboum, with Ethnographical Description*. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.
- Hitchcock, H. R. 1981. *An English-Hawaiian Dictionary with Various Useful Tables*. Tokyo: Charles E. Tuttle Co.
- Hu Peizhou. 1991. "Putaoayay dui Aomenhuade yingxiang" ("The influence of the Putaoya language on the Aomen dialect"). *Fangyan* 4:241-242.
- Jacob, Judith M. 1968. *Introduction to Cambodian*. London: Oxford University Press.
- Jensen, John Thayer. 1977. *Yapese-English Dictionary*. Honolulu: University Press of Hawaii.
- Kamioka Koji and Yamada Minoru. 1979. *Larestani Studies 1: Lari Basic Vocabulary*. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.
- Kawamura Kasajiro, ed. 1981. *Sanseido's Junior Crown Japanese-English Dictionary*. Tokyo: Sanseido.

- Le Ba Kong. 1978. *Standard Pronouncing English-Vietnamese Dictionary*. Houston: Zieleks Co., Inc.
- Lewis, Henry. 1960. *Collins-Spurrell Welsh Dictionary*. London: William Collins Sons & Co., Ltd.
- Li Rulong and Chen Zhanqiang. 1985. "Primary distinctions among the Min dialects." *Chinese Linguistics (Journal of the Linguistic Society of China)* 2:93-173.
- Li Yongming. 1991. "Xinjiapo Chaozhouhua de waiyu jiec he teshu ciyu" ("Foreign loans and peculiar terms in the Chaozhou dialect of Singapore"). *Fangyan* 1:56-63.
- Li Yongsui and Wang Ersong, eds. 1986. *Haniyu Jianzhi (Outline Grammar of the Hani Language)*. Beijing: Minzu Chubanshe.
- Matisoff, James A. 1988. *The Dictionary of Lahu*. University of California Publications in Linguistics, #111. Berkeley and Los Angeles: University of California Press.
- Nakajima Fumio, ed. 1985. *Sanseido's New Concise Japanese-English Dictionary*. Tokyo: Sanseido.
- Nakano Aki'o. 1982. *A Vocabulary of the Beni Amer Dialect of Tigré*. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.
- Panov, Irina. 1985. *New Pocket Romanian Dictionary: Romanian-English, English-Romanian*. New York: Hippocrene Books, Inc.
- Perrot, D. V. 1985. *Teach Yourself Swahili Dictionary*. New York: David McKay Co., Inc.
- Rao Bingcai, Ouyang Jueya, and Zhou Wuji. *Guangzhouhua Fangyan Cidian (Dictionary of the Guangzhou [Cantonese] Dialect)*. Hong Kong: Commercial Press.
- Rashkevich, J., M. Sosar, and L. Timenchik. 1985. *Anglu-Latviešu Vārdnīca (English-Latvian Dictionary)*. Riga: Avots Publishers.
- Rehg, Kenneth, and Damian G. Sohl. 1979. *Ponapean-English Dictionary*. Honolulu: University Press of Hawaii.

- Ruhlen, Merritt. 1987. *A Guide to the World's Languages, Volume 1: Classification*. Stanford: Stanford University Press.
- Saxton, Dean, and Lucille Saxton. 1969. *Dictionary of Papago and Pima to English (O'Odham-Mü-gahn), English to Papago and Pima (Mü-gahn-O'odham)*. Tucson: University of Arizona Press.
- Vanoverbergh, Morice. 1972. *Isneg-English Vocabulary*. Honolulu: University Press of Hawaii.
- Watkins, Calvert. 1985. *The American Heritage Dictionary of Indo-European Roots*. Boston: Houghton Mifflin Company.
- Wolff, Hans. 1969. *A Comparative Vocabulary of Abuan Dialects*. Evanston: Northwestern University Press.
- Wyatt, David K. 1984. *Thailand: a Short History*. New Haven: Yale University Press.
- Yacoubian, Adour H. [n.d.] *English-Armenian and Armenian-English Dictionary, Romanized*. Brooklyn: P. Shalom Publishers, Inc.
- Yap, Elsa P., and Maria V. R. Bunye. 1971. *Cebuano-Visayan Dictionary*. Honolulu: University of Hawaii Press.
- Yuan Jiahua. 1983. *Hanyu Fangyan Gaiyao (Sketch of Chinese Dialects)*. Beijing: Wenzhi Gaige Chubanshe.
- Yukawa Yasutoshi. 1983. "A tentative tonal analysis of Sukuma nouns." *Journal of Asian and African Studies* 26:243-285.
- . 1984. *A Classified Dictionary of the Sambia Language*. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.
- Zhang Zhenxing. 1983. *Taiwan Minnan Fangyan Jilüe (Sketch of the Southern Mtn Dialect(s) of Taiwan)*. Fuzhou: Fujian Renmin Chubanshe.